

ABSTRACT

Methods and apparatus for document based ambiguous character resolution. An application searches a document for words that do not contain ambiguous characters and adds them to a dictionary, then searches the document for words that do contain ambiguous characters. For each ambiguous word, a set of candidate solutions is created by resolving the ambiguous characters in all possible ways. The dictionary is searched for words matching members of the candidate solution set. When a single member is matched, the ambiguous characters are resolved accordingly. When no member or more than one member is matched, a user is prompted to resolve the ambiguous characters. Alternatively, when more than one member is matched, the ambiguous characters are resolved to obtain the largest word, the smallest word, the most words, or the fewest words.

07844-437001